

Measuring clinically Important Change
with the Patient-rated Tennis Elbow Evaluation (PRTEE)

Short title: Measuring change with the PRTEE

Leon Poltawski, Research Fellow, Peninsula College of Medicine & Dentistry, Salmon Pool Lane,
Exeter EX2 8GW. Email: Leon.Poltawski@pcmd.ac.uk

Tim Watson, Professor of Physiotherapy, University of Hertfordshire, Hatfield, AL10 9AB

The authors declare no financial or other interests in the conduct and reporting of this study, which
was funded by the University of Hertfordshire

Keywords:

Tennis elbow; outcome measure; clinical significance.

Abstract

Introduction: The Patient-rated Tennis Elbow Evaluation enables quantitative rating by the patient of pain and functional impairment associated with tennis elbow, or lateral elbow tendinopathy. When used as an outcome measure in trials of therapies, a Minimum Clinically Important Difference (MCID) value is required to interpret trial outcomes. This study aimed to calculate the MCID for a sample of patients diagnosed with lateral elbow tendinopathy (LET).

Methods: The PRTEE was used as an outcome measure with participants in a trial of a novel therapy for LET. It was administered at baseline and after treatment, three weeks later. Score changes were compared to patient-rated Global Change Scores using Receiver Operating Curve analysis. MCID values were calculated for two different criteria of clinically important difference and the effects of baseline symptom severity on the MCID were investigated.

Results: Data was available from fifty seven participants, with PRTEE scores in the range 13-81 / 100. For clinical significance defined as “a little better” the MCID for the total PRTEE score was 7/100 or 22% of baseline score. For clinical significance defined as “much better” or “completely recovered”, the MCID was 11/100 or 37% of baseline score. The MCID value was higher for a subgroup with greater baseline severity.

Conclusions: Substantial changes in the PRTEE scores are required before they can be considered clinically significant. Clinically significant change varies according to the baseline score. The instrument may be less sensitive to change when used by people who are symptomatic in their non-dominant arm.

Introduction

Tennis elbow, also known as lateral elbow tendinopathy or lateral epicondylalgia¹, is a common musculoskeletal disorder that can cause significant pain and disability. Various assessment instruments are used by clinicians and researchers for monitoring the progress of the disorder and the effectiveness of its treatment. A combination of measures is commonly employed, addressing physical variables such as pain and strength, functional limitations and psychosocial factors. The Patient-rated Tennis Elbow Evaluation (PRTEE) is an instrument that has been developed specifically for use with this disorder², and is increasingly being employed in research³⁻⁸. It takes the form of a 15-item questionnaire, with five items addressing pain and ten concerned with functional deficit. For each item, the respondent uses a 0 – 10 numerical scale to rate the average pain or difficulty they have experienced over the previous week while carrying out various activities that are commonly painful in tennis elbow. The marking system ensures pain and function are weighted equally in the total score. Higher scores represent greater severity and the maximum score is 100⁹. Its particular strengths are its simplicity and shortness, and its specificity to tennis elbow. Prior to its development, the only available patient-rated instruments were more generic, such as the Disabilities of the Arm, Shoulder and Hand (DASH) Questionnaire¹⁰ and variants of the SF-36 Quality of Life Questionnaire¹¹.

The measurement properties of the PRTEE have been evaluated in several studies. It has been validated by findings of moderate to excellent correlation with more established scales such as the DASH and SF-36¹²⁻¹⁶, although a recent study found only poor to moderate correlation between individual items in the PRTEE and other comparable measures, including visual analogue pain ratings and items on the EuroQol Scale¹⁷. Several studies using English and other language versions of the PRTEE with people with tennis elbow have concluded that its test-retest reliability is excellent, with correlation coefficients for the total score and the pain and function sub-scores being greater than 0.9 in most cases^{12-16, 18}. Unfortunately the credibility of several of these studies is undermined by

the use of short test-retest intervals - between 30 minutes and two days. Such intervals run the risk of introducing bias, since respondents are likely to remember their previous scores. Calculated reliability indices may thereby be inflated. A more recent study used a four-week test-retest period and calculated Intraclass Correlation Coefficient values of 0.81 for the pain sub-score and 0.76 for the total score¹⁷, still reasonable although not excellent. Such values are probably more reflective of reliability in clinical trials using the PRTEE, where inter-test intervals of several weeks are common. The responsiveness of the PRTEE – its capacity to detect change – has received little attention in the literature, although it is a key property in interpreting score changes¹⁹.

Two dimensions of responsiveness are of interest. The Minimum Detectable Change (MDC) represents the smallest change in score that likely reflects true change rather than measurement error alone²⁰. Knowledge of this quantity is helpful in establishing sample sizes required for clinical trials. It can be calculated from reliability data²¹, but has not been reported in published PRTEE reliability studies. The Minimum Clinically Important Difference (MCID) is the minimum change that would be interpreted as meaningful and worthwhile²². Knowing the MCID value for a particular instrument helps both clinicians and researchers to judge whether treatment is having a useful effect on the variables of interest. In the evaluation of patient-rated measures like the PRTEE, it is the patient's opinion that should be given most weight in defining "meaningful and worthwhile"²³. There are various approaches to the determination of MCID values but an anchor based-method, involving comparison of scores on the measure of interest with those provided by a different but related measure, has been recommended^{23, 24}. Global change ratings are commonly adopted as the anchor referent because they can readily be used to define clinical significance, but this approach has not previously been employed to determine the responsiveness of the PRTEE. Therefore we conducted a study with the aim of determining the MCID of the PRTEE when used as an outcome measure with a sample of individuals with tennis elbow. Since MCID values may vary according to

baseline status on the variables of interest, we also investigated whether this was the case when using the PRTEE.

Materials and Methods

This study formed part of a prospective trial of microcurrent, a subsensory electric current applied to promote analgesia and tissue healing. The trial involved the application of four different forms of microcurrent to compare their effectiveness (manuscript in preparation), and was approved by the authors' Institutional Ethics Committee. Participants were recruited via publicity in local offices, nearby sports clubs and local media, and all provided informed consent before enrolment. Eligible participants were over 18 years of age, able to complete the questionnaire, experiencing lateral elbow pain for at least 3 months with no significant change in the previous month, and had a clinical diagnosis of tennis elbow made by a physiotherapist. The diagnosis was based on a history of lateral elbow pain exacerbated by gripping activities, and tenderness at the lateral epicondyle coupled with lateral elbow pain on at least one of the following: resisted middle finger extension, resisted wrist extension or the chair-lift test^{4, 25}. This was supplemented by sonographic assessment of the common extensor tendon, conducted by the physiotherapist who had been trained for the purpose²⁶. Tendinopathy was confirmed by the presence of thickening, hypoechoic areas, fibrillar disruption or calcification in the tendon^{27, 28}. Concomitant upper limb disorders did not exclude participation so long as the clinical and sonographic signs of tennis elbow were present.

Participants were instructed in the use of the PRTEE following the guidance provided in the user manual⁹. Minor changes were made to the wording of some questions to enhance comprehensibility in the British context. In section 2, the words "coffee" and "milk" were removed, "pants" was replaced by "trousers", and "washcloth or wet towel" were replaced by "wet cloth". For the purposes of this study, participants were asked to complete the questionnaire twice, three weeks apart, during which time they received treatment for their tennis elbow. The treatment consisted of

daily application of microcurrent to the lateral elbow via adherent electrodes. At second assessment participants were also asked to rate any change in their condition since first assessment, using a six point global change score (GCS), with the terms “much worse”, “a little worse”, “unchanged”, “a little better”, “much better”, “completely recovered” as its descriptors. For analytical purposes, these were converted to scores of -2, -1, 0, 1, 2 and 3 respectively. Demographic and medical history data were also collected at baseline.

Data analysis

Absolute and percentage changes in PRTEE total and subscale scores between the assessments were calculated. These were compared with corresponding GCS values graphically and using the Spearman rank correlation coefficient, since the GCS is an ordinal scale. This was done to establish whether scores on the two scales were related, which is necessary if the GCS is to be used as the external comparator for the PRTEE scores²³. The MCID for the instrument was then determined using the GCS as an anchor and determining the group average change in PRTEE score corresponding to two different definitions of clinical significance: CGS = 1 (“a little better”) and GCS = 2 (“much better”). The specificity and sensitivity of the change score for these classifications were examined using a Receiver Operator Characteristic (ROC) curve analysis, and the MCID was estimated for the best match of sensitivity and specificity²⁹. Area under the curve (AUC) analysis was conducted to gauge how well the PRTEE discriminates between those whose score change is clinically significant from those for whom it is not. To investigate the potential effect of baseline severity on the MCID, separate analyses were conducted on subgroups defined by a baseline total PRTEE score of <40 and ≥ 40. This cut-off was selected post-hoc to ensure reasonably sized sub-groups for this sample. All statistical tests were conducted using SPSS 17 (SPSS Inc, Chicago) with significance set at $p \leq 0.05$.

Results

Sixty people diagnosed with tennis elbow underwent baseline assessment. Of these, one withdrew before second assessment and two had missing questionnaires due to administrative error, leaving data from 57 people for analysis. Their baseline characteristics are presented in Table 1. PRTEE subscale and total scores covered a wide range of the possible score values. Eleven participants were symptomatic in their non-dominant arm. Five of these reported bilateral symptoms and had baseline total scores in the range 34-76/100, representing a range of severities.

At second assessment, global change scores indicated that three participants judged their condition to have deteriorated, 18 said it was unchanged, 19 reported it was a little better and 17 said it was much better. Chart 1 is a box and whisker plot showing the median and inter-quartile range of change in PRTEE total scores corresponding to each global change score. The scores of the two outliers were checked and found correct. As the plot illustrates, the spread of score changes increased with higher GCS values; however, the spread of percentage score changes varied much less between GCS scores. Absolute and percentage changes in pain, function and total scores all showed significant ($p < 0.001$) moderate correlation with global change scores, with Spearman's rho between -0.54 and -0.66, the highest values being for the total score changes. The negative correlations are a consequence of the way score changes are calculated: by subtraction of baseline from follow-up values. Clinical improvements therefore corresponded to negative change values. Changes in pain and function subscale scores were significantly correlated (Spearman's rho = 0.63, $p < 0.01$).

Table 2 presents data from the ROC curve analysis, conducted using the two definitions of clinically significant change and estimating a cut-off value for each score where sensitivity and specificity are approximately equal.

AUC values greater than 0.8 are considered to indicate excellent discriminatory capacity²², and all scale and sub-scale changes met this criterion except for percentage changes in the function subscale score, which approached it. As would be expected, larger absolute and percentage changes in scores are required to meet the stricter criterion for the MCID, with score changes of 35-40% of the baseline value being required to be considered clinically significant. Table 3 presents data from the sub-group analysis, with cut-off values for total score where clinical significance is defined as GCS=2. Both absolute and percentage changes in total score had AUC values greater than 0.8.

A number of participants volunteered comments about use of the questionnaire. A common comment is that it was problematic to score “difficulty” and “pain” as separate entities because the difficulty was seen as entirely a consequence of pain rather than, say, weakness; so they were inclined to see the two subscales as addressing the same question. Some also reported that, since they were either symptomatic on the non-dominant side or had learned to use the non-symptomatic limb for activities identified in the questionnaire, they did not feel limited in the activities specified.

Discussion

Since pain and functional limitations are the most common consequences of tennis elbow, it is reasonable to expect changes in their PRTEE scores to correlate with the individual’s overall rating of the condition. A correlation threshold of 0.30 – 0.35 between a patient-rated outcome and an anchor has been recommended for estimating MCIDs, and this criterion was met in this study²³. However, the correlation is only moderate, suggesting that factors other than those addressed by the questionnaire are involved in the patient’s global impression. This interpretation is supported by the relatively large changes in PRTEE scores that are required before the individual considers that significant improvement has occurred. It may also be relevant that more than 20% of participants in this study either were symptomatic in the non-dominant arm, or reported compensatory use of the unaffected arm. A consequence of this is that changes in the condition may not have been

accompanied by commensurate changes in rating of those particular items by these individuals, in which case the instrument is less sensitive to change for this sub-group. The five participants who reported bilateral symptoms had a variety of global change scores and inspection of their data suggested that these were commensurate with their PRTEE score changes. Hence, the questionnaire appeared valid for use with this sub-group.

The difficulty that was reported by some participants in distinguishing ratings of pain and difficulty attempting tasks is reflected in the significant correlation between changes in the pain and function subscale scores. Pain and function scores were also significantly correlated at each time point. This suggests that respondents largely equated the two constructs and casts some doubt on the value of modelling them as distinct metrics in a patient-rated measure. It may be that separate rating of functional problems is worthwhile when weakness is a significant feature of the presentation.

The primary function of the MCID is to inform interpretation of group mean values, for example in clinical trials of treatment effectiveness. This study indicates that substantial changes in group mean PRTEE scores are required before a clinically significant change can be assumed to have occurred, even using the more liberal criterion of “a little better”. The more stringent criterion of “much better” requires falls of the order of 35-40% of baseline scores to be confident that a desirable level of improvement has occurred. Where a study reports a statistically significant difference between two treatments, but the actual change in the more effective treatment group mean PRTEE score is less than 35%, the value of the treatment remains open to question.

The sub-group analysis confirms that the absolute score change considered clinically significant depends on the baseline severity of the condition. Sub-groups with milder symptoms require considerably smaller PRTEE score changes than those with more severe presentations in order to consider that significant improvements have occurred. However, percentage changes in scores that

are regarded as clinically significant are similar between severity sub-groups. This finding is consistent with studies of numerical pain scales, which have concluded that changes of 30-35% of baseline score are clinically significant^{30, 31}.

The Minimum Detectable Difference can be determined using test-retest reliability data²¹, although most reliability studies have not reported it for the PRTEE. Calculations using data from three studies^{13, 16, 18} suggests that the MDD for the total score lies in the range 8-12 points, meaning that the score must change by at least this much to be confidently interpreted as real, and not attributable to random error alone. These values exceed the liberal MCID value for milder presentations, and so suggest that the test-retest reliability of the instrument may limit its use for detecting clinically significant change in this group.

There were several potential limitations in this study. The changes made to the wording of some items in the questionnaire may have influenced its measurement properties. Indeed this was the intention: we felt that the relatively minor changes would enhance understanding and consistent interpretation of the questions. This might have produced different findings than if the standard version had been used, although we suspect that the differences would be minor. The study sample size was not determined prospectively because the investigation drew data from a broader pilot study. This was a pragmatic pilot trial, whose liberal inclusion criteria mean that concomitant upper limb disorders may have contributed to the elbow pain, and so influenced PRTEE scores. In this sense, the questionnaire was being applied to a broader population than its name suggests. However, we would argue that since tennis elbow often presents clinically with other upper quadrant problems, testing the questionnaire with this broader population assesses its value in a realistic context. Several other studies evaluating the measurement properties of the PRTEE have not excluded participants with common upper limb co-morbidities such as radiculopathy or radial

nerve involvement^{13, 15, 32}, and it has also been used with different groups altogether, for instance those undergoing arthroplasty³³.

Objections have been raised regarding the retrospective use of global change measurements as anchors for determining responsiveness³⁴: recall bias may mean that change scores correlate with present status, and the validity and reliability of a subjectively-rated GCS is very difficult to evaluate. Empirically derived data in a variety of clinical contexts suggests that the method does have external validity³⁵, but comparisons with MCID values derived using other methods are required to build confidence in the interpretation of PRTEE score changes.

It has been argued that responsiveness studies should be prospective, based on reasonable assumptions about whether and what kind of change is expected in the sample³⁶. Although this analysis was retrospective, it used data from a study in which such differences were expected: the hypothesis of the trial, based upon previous studies, was that some types of microcurrent would be more effective than others. Nevertheless the results of this analysis require corroboration by other, prospectively designed studies.

Conclusions

This study has provided estimates of the changes in PRTEE scores that are required to judge whether significant changes have occurred in symptoms experienced by patients with tennis elbow. For a group using the questionnaire, mean PRTEE score reductions of at least 11 points - or an improvement of 37% on the mean baseline score - is necessary to consider that a substantial improvement has taken place. Using a less stringent criterion, falls of 7 points or 22% of the baseline score can be interpreted as indicating a limited but meaningful improvement. The value of separate ratings for pain and function in this instrument, and used with this population, is open to question. Until this matter is resolved, the total score may be of value in estimating symptom severity. The

questionnaire may have lower sensitivity to change when used with respondents who are symptomatic in the non-dominant arm.

1. Stasinopoulos, D. and M.I. Johnson, 'Lateral elbow tendinopathy' is the most appropriate diagnostic term for the condition commonly referred-to as lateral epicondylitis. *Med Hypotheses*, 2006. **67**(6): p. 1400-2.
2. MacDermid, J., Update: The Patient-rated Forearm Evaluation Questionnaire is now the Patient-rated Tennis Elbow Evaluation. *J Hand Ther*, 2005. **18**(4): p. 407-10.
3. Faes, M., B. van den Akker, J.A. de Lint, *et al.*, Dynamic extensor brace for lateral epicondylitis. *Clin Orthop Relat Res*, 2006. **442**: p. 149-57.
4. Martinez-Silvestrini, J.A., K.L. Newcomer, R.E. Gay, *et al.*, Chronic lateral epicondylitis: comparative effectiveness of a home exercise program including stretching alone versus stretching supplemented with eccentric or concentric strengthening. *J Hand Ther*, 2005. **18**(4): p. 411-9, quiz 420.
5. Nilsson, P., E. Thom, A. Baigi, *et al.*, A prospective pilot study of a multidisciplinary home training programme for lateral epicondylitis. *Musculoskeletal Care*, 2007. **5**(1): p. 36-50.
6. D'Vaz, A.P., A.J. Ostor, C.A. Speed, *et al.*, Pulsed low-intensity ultrasound therapy for chronic lateral epicondylitis: a randomized controlled trial. *Rheumatology (Oxford)*, 2006. **45**(5): p. 566-70.
7. Coombes, B.K., L. Bisset, L. Connelly, *et al.*, Optimising corticosteroid injection for lateral epicondylalgia with the addition of physiotherapy: A protocol for a randomised control trial with placebo comparison. *BMC Musculoskelet Disord*, 2009. **10**(1): p. 76.
8. Alizadehkhayat, O., A.C. Fisher, G.J. Kemp, and S.P. Frostick, Pain, functional disability, and psychologic status in tennis elbow. *Clin J Pain*, 2007. **23**(6): p. 482-9.
9. MacDermid, J., The Patient-Rated Tennis Elbow Evaluation (PRTEE) User Manual. 2007, School of Rehabilitation Science, McMaster University: Hamilton, Canada.
10. Hudak, P.L., P.C. Amadio, and C. Bombardier, Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). *Am J Ind Med*, 1996. **29**(6): p. 602-8.

11. Ware, J.E., Jr. and C.D. Sherbourne, The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*, 1992. **30**(6): p. 473-83.
12. Altan, L., I. Ercan, and S. Konur, Reliability and validity of Turkish version of the patient rated tennis elbow evaluation. *Rheumatol Int*, 2009.
13. Nilsson, P., A. Baigi, B. Marklund, and J. Mansson, Cross-cultural adaptation and determination of the reliability and validity of PRTEE-S (Patientskattad Utvardering av Tennisarmbage), a questionnaire for patients with lateral epicondylalgia, in a Swedish population. *BMC Musculoskelet Disord*, 2008. **9**(1): p. 79.
14. Newcomer, K.L., J.A. Martinez-Silvestrini, M.P. Schaefer, *et al.*, Sensitivity of the Patient-rated Forearm Evaluation Questionnaire in lateral epicondylitis. *J Hand Ther*, 2005. **18**(4): p. 400-6.
15. Rompe, J.D., T.J. Overend, and J.C. MacDermid, Validation of the Patient-rated Tennis Elbow Evaluation Questionnaire. *J Hand Ther*, 2007. **20**(1): p. 3-10; quiz 11.
16. Leung, H.B., C.H. Yen, and P.Y. Tse, Reliability of Hong Kong Chinese version of the Patient-rated Forearm Evaluation Questionnaire for lateral epicondylitis. *Hong Kong Med J*, 2004. **10**(3): p. 172-7.
17. Chung, B. and J.P. Wiley, Validity, responsiveness and reliability of the Patient-Rated Tennis Elbow Evaluation. *Hand Therap*, 2010. **15**(3): p. 62-68.
18. Overend, T.J., J.L. Wuori-Fearn, J.F. Kramer, and J.C. MacDermid, Reliability of a patient-rated forearm evaluation questionnaire for patients with lateral epicondylitis. *J Hand Ther*, 1999. **12**(1): p. 31-7.
19. Guyatt, G., S. Walter, and G. Norman, Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis*, 1987. **40**(2): p. 171-8.
20. Stratford, P.W., F.M. Binkley, and D.L. Riddle, Health status measures: strategies and analytic methods for assessing change scores. *Phys Ther*, 1996. **76**(10): p. 1109-23.
21. Weir, J.P., Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*, 2005. **19**(1): p. 231-40.

22. Copay, A.G., B.R. Subach, S.D. Glassman, *et al.*, Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J*, 2007. **7**(5): p. 541-6.
23. Revicki, D., R.D. Hays, D. Cella, and J. Sloan, Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*, 2008. **61**(2): p. 102-9.
24. de Vet, H.C., C.B. Terwee, R.W. Ostelo, *et al.*, Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes*, 2006. **4**: p. 54.
25. Haker, E., Lateral epicondylalgia: Diagnosis, treatment and evaluation. *Crit Rev Phys Rehabil Med*, 1993. **5**(2): p. 129-154.
26. Poltawski, L., S. Ali, V. Jayaram, and T. Watson, Reliability of sonographic assessment of tendinopathy in tennis elbow. *Skeletal Radiol*, 2011.
27. Poltawski, L., V. Jayaram, and T. Watson, Measurement issues in the sonographic assessment of tennis elbow. *J Clin Ultrasound*, 2010. **38**(4): p. 196-204.
28. Connell, D., F. Burke, P. Coombes, *et al.*, Sonographic examination of lateral epicondylitis. *AJR Am J Roentgenol*, 2001. **176**(3): p. 777-82.
29. Farrar, J.T., R.K. Portenoy, J.A. Berlin, *et al.*, Defining the clinically important difference in pain outcome measures. *Pain*, 2000. **88**(3): p. 287-94.
30. Salaffi, F., A. Stancati, C.A. Silvestri, *et al.*, Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain*, 2004. **8**(4): p. 283-91.
31. Farrar, J.T., J.P. Young, Jr., L. LaMoreaux, *et al.*, Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*, 2001. **94**(2): p. 149-58.
32. Chung, B. and J.P. Wiley, Validity, responsiveness and reliability of the Patient-Rated Tennis Elbow Questionnaire *Hand Therap*, 2010. **15**: p. 62-68.
33. Angst, F., M. John, G. Pap, *et al.*, Comprehensive assessment of clinical outcome and quality of life after total elbow arthroplasty. *Arthritis Rheum*, 2005. **53**(1): p. 73-82.

34. Norman, G.R., P. Stratford, and G. Regehr, Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. *J Clin Epidemiol*, 1997. **50**(8): p. 869-79.
35. Osaba, D. and M. King, Meaningful Differences, in *Assessing Quality of Life in Clinical Trials*, P. Fayers and R. Hays, Editors. 2005, Oxford University Press: Oxford. p. 243-257.
36. Stratford, P.W. and D.L. Riddle, Assessing sensitivity to change: choosing the appropriate change coefficient. *Health Qual Life Outcomes*, 2005. **3**: p. 23.

Table 1: baseline characteristics of participants.

Sex	30 male, 27 female
Age - mean (range)	53 (40-69) years
Arm dominance	10 left, 46 right, 1 ambidextrous
Symptomatic arm	13 left, 39 right, 5 bilateral
Number with dominant arm symptomatic	46 (80%)
Symptom duration – median (range)	13 (3-240) months
PRTEE Pain sub-scale – mean (range)	21 (6-48) / 50
PRTEE Function sub-scale – mean (range)	18 (2-42) / 50
PRTEE Total score – mean (range)	39 (13-81) / 100

Table 2: AUC analysis and cut-off values for PRTEE score changes

	Clinical significance set at GCS ≥ 1		Clinical significance set at GCS ≥ 2	
	Area under curve	Cut-off value	Area under curve	Cut-off value
	(95% CI)		(95% CI)	
Absolute change in pain score / 50	0.84 (0.74 -0.94)	4	0.8 (0.68-0.92)	6
Percentage change in pain score	0.81 (0.70-0.93)	20%	0.82 (0.70-0.94)	35%
Absolute change in function score / 50	0.83 (0.72-0.93)	3	0.83 (0.73-0.94)	6
Percentage change in function score	0.78 (0.65-0.91)	25%	0.81 (0.69-0.93)	40%

Absolute change in total score / 100	0.86 (0.77-0.95)	7	0.83 (0.72-0.94)	11
Percentage change in total score	0.83 (0.72-0.93)	22%	0.86 (0.73-0.96)	37%

Table 3: MCID values for subgroups with different PRTEE total scores at baseline

	Cut-off value	
	Baseline total score < 40/100	Baseline total score ≥ 40/100
Absolute change in total score / 100	7	21
Percentage change in total score	35%	40%

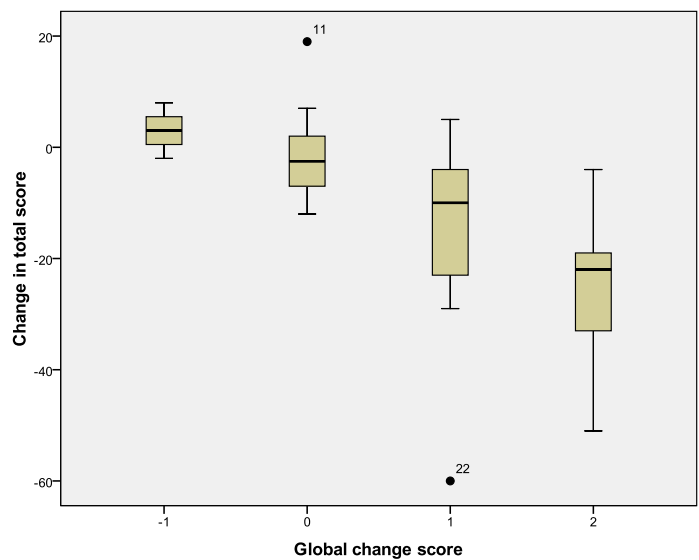


Figure 1: Change in PRTEE total score between assessments for each Global Change Score